

SEQUENCE LISTING

SEQ ID NOs:1 and 2

LOCUS COTMYBA 1006 bp mRNA PLN 31-DEC-1993  
DEFINITION Cotton DNA-binding domain mRNA.  
5 ACCESSION L04497  
NID g437326  
VERSION L04497.1 GI:437326  
KEYWORDS  
10 SOURCE Gossypium hirsutum (cultivar Acala SJ-2) 3-day pre-anthesis ovule  
cDNA to mRNA.  
ORGANISM Gossypium hirsutum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
15 REFERENCE 1 (bases 1 to 1006)  
AUTHORS Wilkins,T.A. and Lu,C.-C.  
JOURNAL Unpublished (1993)  
FEATURES Location/Qualifiers  
source 1..1006  
20 /organism="Gossypium hirsutum"  
/cultivar="Acala SJ-2"  
/db\_xref="taxon:3635"  
/dev\_stage="3-day pre-anthesis"  
/tissue\_type="ovule"  
mRNA 1..1006  
25 CDS 59..943  
/note="MYB A; putative"  
/codon\_start=1  
/protein\_id="AAA33067.1"  
/db\_xref="PID:g437327"  
/db\_xref="GI:437327"  
30  
" protein\_bind 92..403  
/note="putative"  
/bound\_moiety="MYB"  
repeat\_region 107..133  
/note="MYB DNA-binding domain repeat signature 1;  
putative"  
misc\_feature 323..394  
40 BASE COUNT 323 a 209 c 204 g 270 t  
ORIGIN  
1 taacaccgtt attctttctc tattctacctt gatttgattt tgtaactgtat  
61 gggacgatca ctttgttgtaaaaggctca taccacaaaa ggtgcctgga ccaaagagga  
121 agatcaacgc ctcataact acatccgtt ccattggtaa ggctgctggc gttccccc  
45 181 caaaagctgctt gggctgttta gatgtggtaa gagttcaga ttaagatgaa taaactactt  
241 gaggcctgtat cttaaaggagag gaaatttcac tgaagaagaa gatgagctt acatcaagct  
301 tcacagttta ctggaaaca aatggtcatt gattgtggaa agattaccag gaagaacaga  
361 taatgagata aagaactactt ggaacacaca catcaaaaaga aagcttataa gcagaggaat  
421 tgatccacaa actcatcgat ctctcaatca aacggccat accaacacag tcacagcccc

481 caccgaattg gatttcagaa actcgccac atccgttcc aaatccagtt ccatcaaaaa  
541 cccgtctcg gattcaatt acaatgaatt tcaattcaag tccaacacag attcccttga  
601 agaaccac acgtacggca gcagtggcat gactacagat gaagagcaac aagaacagct  
661 gcacaagaag cagcaatacg gtccgagcaa tggcaagac ataaatttg agctgtcgat  
721 tggattgtt tcagctgact catctcggtt atcaaattgc aactcgccg agtcgaaacc  
781 aaaggttagat aacaacaatt tccagttct tgaacaagct atgggtggcta aggccgtatg  
841 ttgtgttgg caatttaggtt ttgaaacaag tggaaattgtt aggaactgtc aaaattcaaa  
901 ttcaaattggc ttctatagtt attgttagacc ctggattca tagggtcatc ttttcttct  
961 ttcttcgtt ttttaggaga taaaattat cttaaattt aaaaaaa

10

MGRSPCCEKAHTNKGAWTKEEDQRLINYIRVHGEGCWRSLPKAALLRCGKSCRLR  
WINYLRPDLKRGNFTEEEDELIKLHSLLGNKWSLIAGRLPGRTDNEIKNYWNTHIKR  
KLISRGIDPQTHRPLNQTANTNTVAPTELDFRNSPTSVKSSEIKNPSLDFNYNEFQF  
15 KSNTDSLEEPNCTASSGMTTDEEQQEQLHKKQQYGPSNGQDINLELSIGIVSADSSRV  
SNANSAESKPKVVDNNNFQFLEQAMVAKAVCLCWQLGFTSEICRNCQNSNSNGFYS  
YCRPLDS

20 **SEQ ID NO:3 and 4**

LOCUS AF034134 1151 bp mRNA PLN 02-MAR-1998  
DEFINITION Gossypium hirsutum MYB-like DNA-binding domain protein (Cmy-O)  
mRNA, complete cds.

ACCESSION AF034134

25 NID g2921339

VERSION AF034134.1 GI:2921339

KEYWORDS

SOURCE upland cotton.

ORGANISM Gossypium hirsutum

30 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicots; core  
eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 1151)

AUTHORS Loguercio,L.L., Zhang,J. and Wilkins,T.A.

35 TITLE Structure and expression of six classes of myb-domain genes in  
allo tetraploid cotton (Gossypium hirsutum L.)

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1151)

AUTHORS Loguercio,L.L., Zhang,J. and Wilkins,T.A.

40 TITLE Direct Submission

JOURNAL Submitted (13-NOV-1997) Agronomy & Range Science, University of  
California, One Shields Ave., Davis, CA 95616-8515, USA

FEATURES Location/Qualifiers

source 1..1151

45 /organism="Gossypium hirsutum"  
/cultivar="Acala SJ-2"  
/db\_xref="taxon:3635"  
/dev\_stage="3 days pre-anthesis"  
/tissue\_type="ovule"

gene 1..1151  
/gene="Cmy-O"  
/note="MYB-domain gene O"  
CDS 72..752  
/gene="Cmy-O"  
/function="putative MYB-like transcription factor"  
/note="similar to MYB A encoded by GenBank Accession  
Number L04497"  
/codon\_start=1  
/product="MYB-like DNA-binding domain protein"  
/protein\_id="AAC04720.1"  
/db\_xref="PID:g2921340"  
/db\_xref="GI:2921340"

15 BASE COUNT 382 a 215 c 240 g 314 t  
ORIGIN

1 cgattttcttccccgtgttggttgcacagaaagttagagaaagtttacttttgatt  
61 ttgaaaactccgatgagaaaaaccttgctcgataaaacaaggccaacaaggggagcctgg  
121 ccaagcaagaagatcaaaagctcattgattatatacgatcatggtaaaaggctgtggc  
181 gttccctccc caaagctgca ggtttgcaccgttgcgttaaagttgcaggctgagatgga  
241 taaattactt aagaccagatatcaaactgtgtaaacttgc tcaagacgaa gaggactaa  
301 ttatcaaactccatgctctcttggtaaccgggttgcactgatagctggtagattaccag  
361 gaagaacaga taatgaagtg aagaactattgaaattccataaaagaga aagctaata  
421 agatggggatcgatcctaataaccataagt tgaaccaatatcctcatcatgttgc  
481 ttaaccccac caccaccaacctcatggatgtggcatgtaa gcttagatgttgtcaacag  
541 acaatgatgatgatggatctca gatgtgcatttatctgaagacgcaaca ccggccactg  
601 gtatccaa ctggacctt gatctcacaatttgctttccatcgagtccatcaagaata  
661 ttatgaaga aagccagcagaaaacagcatctatgtaacaaatgtatgaa gaagaacaat  
721 atacagtccc tacccttcttcttcatgatgagaaaaaaaaggcctcacatgtgg  
781 agattcgtcaaaaagaccta aaggcttacg aaggcaacatgcacgcattgtcaattct  
841 ttggatgatggattgaaac catafccttgcatttagaaaggagaaataagctaaaa  
901 ctgtattattgtgtataaatttggtagaaa gaaagattc aacttaagaaatggatcaa  
961 ataactgaatgaatgaacgatattgcagataagttttaggagggttcaatcaacttac  
1021 tgcaattatatttggtggagctgtatgttaggtatgtagttc atcgatcatg aactgaacct  
1081 ttgatatttcaggcttaatgtctgttttatgcgtaaa gatattctcaatgtgagat  
1141 cagctaaaaaa a

MRKPCCDKQGTNKGAWSKQEDQKLIDYIRIHGEWCWRSLPKAAGLHRCGKSCRLR  
40 WINYLRPDIKRGNFAQDEEDLIKLHALLGNRWLSIAGRLPGRTDNEVKNYWNSHIK  
RKLMKMGIDPNHKLNNQYPHVGPLNPTTNSMDVACKLRVCSTDNDGISDAAS  
YLEDATPPTGISNLDDLTIAFPSSPIKNIEESQQKTASIVTNDEEQYTVPTLLLFR  
repeat\_region 102..257  
/note="putative MYB DNA-binding domain repeat R2"  
repeat\_region 258..412  
/note="putative MYB DNA-binding domain repeat R3"

**SEQ ID NO: 5**

Cotton (*Gossypium hirsutum* L. cv. Acala SJ-2) 10 dpa fiber cDNA

R2R3-MYB transcription factor

GhMYB7

5 ID 6W DNA; 1081 BP.

10 SQ SEQUENCE 1081 BP; 353 A; 228 C; 240 G; 260 T;  
CTCCCCGCGG TGGCGGCCGC TCTAGAACTA GTGGATCCCC CGGGCTGCAG  
GAATTCCGGCA

CGAGGAAAGA AGTGTGAAAA AAAAAATGGG AAGGAGTCCT TGTTGTTCTA  
AGGAAGGCCT

15 TAACAGAGGA GCTTGGACTG CTCTTGAAGA CAAAATTCTT AAAGATTATA  
TCAAAGTACA

CGGTGAAGGT CGTTGGAGAA ATCTCCCCAA AAGAGCTGGT CTTAAGAGAT  
GTGGGAAAAG

TTGTAGGCTT CGGTGGTTGA ATTATTGAG ACCTGATATT AAAAGAGGTA

20 ACATATCACC  
TGACGGAGGAA GAGCTTATCA TCAAACCTCA CAAACTCTG GGAAACAGAT  
GGTCTTGAT

AGCTGGGAGG CTTCCAGGAC GAACAGACAA TGAAATAAAG AATTACTGGA  
ACACCAACTT

25 AAGTAAAAGA GTTCCGATC GTCAAAAGTC ACCCGCCGCT CCTTCGAAAAA  
ATCCCAGGC

GGCTCGACGA GGAACTGCTG GTAATGGCAA TACCAATGGT AATGGTAGTG  
GTAGTTCTC

GACACACGTG GTGCGGACAA GGGCGACAAG GTGCTCCAAG GTTTCATCAA  
ACCCCTCCTCA

30 CTACACACAA AACAGAGACC CAAAGCCTTC TTCAACTTGT TCAAATCATG  
GGGATCACCG

GGAACCTAAA ACAATGAATG AGTTGTTATT ACCGATAATG TCAGAATCCG  
AGAATGAAGG

35 GACGACCGAT CATATATCAT CGGATTTCAC ATTTGACTTC AACATGGGAG  
AGTTTTGTTT

ATCGGATCTT TTGAATTCCG ATTTCTGCGA TGTAAACGAG CTTAATTACA  
GCAATGGTTT

TGATTCGTCA CCCTCACCGG ATCAGCCTCC TATGGATTTC TCCGACGAAA

40 TGCTAAAAGA  
GTGGACGGCC GCCGCCTCCA CTCACTGCTG TCACCAAAGT GCGGCTTCCA  
ATCTCCAGTC

CTTGCCTCCA TTTATTGAAA ATGGAATTGA ATGACCTTGA AAAAAATAAAA  
GACGAAAAAT

45 ATTTCTCAT GTAAACTAAA TAAACACATC TTCCATCATT AAAAAAAAAAA  
AAAAAAAAAA

A

/

**SEQ ID NO:6**

ID 6WP PRT; 302 AA.

DT 09-NOV-1999 (CREATED BY PC/GENE PROGRAM TRANSL)

5 DE

CC TRANSLATED FROM DNA SEQUENCE 6W (BASES 86 TO 993).

SQ SEQUENCE 302 AA; 33766 MW; 451675 CN;

MGRSPCCSKE GLNRGAWTAL EDKILKDYIK VHGEGRWRNL PKRAGLKRCG  
KSCRRLWLN

10 LRPDIKRGN NI SPDEEEELI IK LHKL LGNR WS LIAGRLPGRT DNEIKNYWNT  
NLSKRVSDRQ

KSPAAPSKNP EAARRGTAGN GNTNGNGSGS SSTHVVVRTRA TRCSKVFINP  
PHYTQNRDPK

PSSTCSNHGD HREP KTMNEL LLPIMSESEN EGTTDHIS SD FTFDFNMGEF

15 CLSDLLNSDF CDVNELNYSN GFDSSPSPDQ PPMDFSDEML KEWTAAASTH CCHQSAASNL  
QSLPPFIENG

IE

20

Cotton (*Gossypium hirsutum* L. cv Acala SJ-2) 10 dpa fiber cDNA

R2R3-MYB Transcription Factor

GhMYB8

5   **SEQ ID NO: 7**

ID 19W; DNA; 944 BP.

SQ SEQUENCE 944 BP; 319 A; 163 C; 214 G; 246 T; 2 OTHER;

TGGAGCTCCC CGCGGTGGCN GNCGCTCTAG AACTAGTGGATCCCCCGGGC

TGCAGGAATT

10   CGGCACGAGA CTCCAACAAA TGTCCATGAA AAAAGAAGGT GAAATTCTAT  
ACAAAAAAGGG

ATTATGGGCA ATGGAGGAAG ACAAGTTACT CATTGATTAT GTCAATGTCC  
ATGGAAAAAGG

ACAATGGAAC AAAATAGCCA ACAGAACAGG TTTGAAGAGA AGTGGGAAAA

15   GTTGTGGCT

AAGGTGGATG AATTACCTGA GTCCTAACGT TAAAAAAGGGT GATTTTCTG  
AAGAAGAAGA

AGACCTCGTC ATTAGACTTC ATAAGCTTCT TGGAAACAGG TGGTCTTGAT  
TTGCGAAACG

20   AGTTCCAGGT CGAACTGACA ATCAAGTCAA GAATTACTGG AATAGTCATT  
TGAGGAAGAA

ACTAGGGATC ATTGATCAA ACAAGACAAG GATCGATTT TGTCAAAGTT  
CAAAGCAAGT

CAAAGTGTGT CATGTTGATG AGGCAGGCCAC GGATCCAAGT CCTGGACATG  
GAACAAACCAC

25   TGAAACCACG GGTATAACAG TGGATCAGAG TAACCAGCAG GAAGTCATTG  
ATCATCGGGT

CTTAAACAAAT ACTACTCAAG AATCAATGAC CAGTGAGAGT TATATCAACA  
CTTTCTGGAT

30   TCCTGACCAT GATTATGAGC TAAGTACACT TGCCATGATT GACCATGATT  
ATGAGCTAAG

TACACTTGCC ATGATTGACC ACTTCCATGA ATGTTCTTCT TTTCATCTTA  
GCTAGAGACT

ATGTTATTAG ATTGGGTTT TATTTTAGA TATAAGTATG CATCTAACAT

35   GGCAATGTTA

AATTTTCAA AAGATTTTC ATGTATTGA GCAGTTCATG TGTTGAAGA  
TTAAGATATT

TCTGAAAAAA AAAAAAAA AAAACCGAGG GGGGCCGGT ACCC

//

40   **SEQ ID NO:8**

ID 19WP; PRT; 231 AA.

DT 25-OCT-1999 (CREATED BY PC/GENE PROGRAM TRANSL)

45   DE

CC TRANSLATED FROM DNA SEQUENCE 19W (BASES 80 TO 772).

SQ SEQUENCE 231 AA; 26698 MW; 283450 CN;

MSMKKEGEIL YKKGLWAMEE DKLLIDYVNV HGKGQWNKIA NRTGLKRSGK  
SCRLRWMNYL

SPNVKKGDFS EEEEDLVIRL HKLLGNRWSL IAKRVPGRKD NQVKNYWNSH  
LRKKLGIDIQ  
NKTRIDFCQS SKQVKVCHVD EAATDPSPGH GTTTETTGIT VDQSNQQEVI  
DHRVLNNNTQ  
5 ESMTSESYIN TFWIPDHHDYE LSTLAMIDHD YELSTLAMID HFHECSSFHL S